Sentinel – MSI: – File name conventions and product folder description

SENTINEL-2 products will be made available to users in SENTINEL-SAFE format, including image data in JPEG2000 format, quality indicators (e.g. defective pixels mask), auxiliary data and metadata.

The SAFE format has been designed to act as a common format for archiving and conveying data within ESA Earth Observation archiving facilities. The SAFE format wraps a folder containing image data in a binary data format and product metadata in XML. This flexibility allows the format to be scalable enough to represent all levels of SENTINEL products.

A SENTINEL-2 product refers to a directory folder that contains a collection of information (Figure 1). It includes:

- a *manifest.safe* file which holds the general product information in XML
- a preview image in JPEG2000 format
- subfolders for measurement datasets including image data (granules/tiles) in GML-JPEG2000 format
- subfolders for datastrip level information
- a subfolder with auxiliary data (e.g. International Earth Rotation & Reference Systems (IERS) bulletin)
- HTML previews

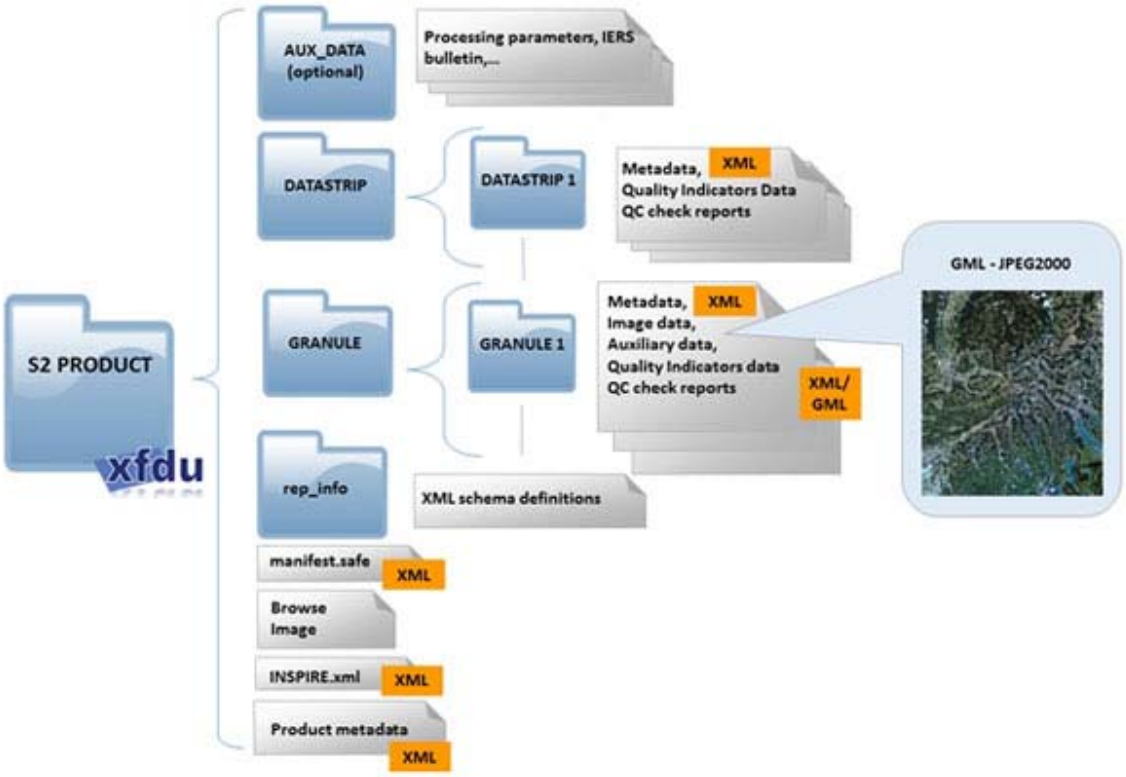

Figure 1: SENTINEL-2 product physical format

LEVEL-2A Data Format

The Level-2A prototype product is an orthorectified product providing Bottom-Of-Atmosphere (BOA) reflectances, and basic pixel classification (including classes for different types of cloud). The generation of this prototype product is carried out by the User from Level-1C products.

The Level-2A image data product uses the same tiling, encoding and filing structure as Level-1C.

The Level-2 product is also in SAFE format, which groups together several types of file:

- metadata file (XML file)
- preview image (JPEG2000 with GML geo-location)
- tiles files with BOA reflectances image data file (GML / JPEG2000) for each tile
- datastrip files
- auxiliary data
- ancillary data (Ground Image Processing Parameters (GIPPs))

More information on Level-2A format can be found in the SENTINEL-2 Technical Guide.

New format Naming Convention for Sentinel-2 Level-1C products generated after the 6th of December, 2016:

As of the above date, all newly acquired (NRT) Level-1C products from the DataHub are distributed in a new product format [(1)]. This format has been introduced in order to overcome the 256 characters limitation on pathnames imposed by Windows platforms. This has been achieved by compaction of the filenames, including the naming of internal folders and files.

The updated product format is described in detail in the Sentinel-2 Products Specification Document (PSD) (version 14.2). The product format corresponding to the new Level-1C product combines the features described in the 'Compact Naming Convention' and the 'Complete Single Tile' sections in the PSD.

NOTE: In updating the content of this page, an error has been spotted in the PSD (Section 4.2.1.1 Product Name Root Directory). The information given below is correct.

Compact Naming Convention

The compact naming convention is arranged as follows:

MMM_MSIL1C_YYYYMMDDHHMMSS_Nxxyy_ROOO_Txxxxx_<Product Discriminator>.SAFE

The products contain two dates.

The first date (YYYYMMDDHHMMSS) is the datatake sensing time.
The second date is the "<Product Discriminator>" field, which is 15 characters in length, and is used to distinguish between different end user products from the same datatake. Depending on the instance, the time in this field can be earlier or slightly later than the datatake sensing time.

The other components of the filename are:

MMM: is the mission ID(S2A/S2B)
MSIL1C: denotes the Level-1C product level
YYYYMMDDHHMMSS: the datatake sensing start time
Nxxyy: the Processing Baseline number (e.g. N0204)
ROOO: Relative Orbit number (R001 - R143)
Txxxxx: Tile Number field
SAFE: Product Format (Standard Archive Format for Europe)

Thus, the following filename S2A_MSIL1C_20170105T013442_N0204_R031_T53NMJ_20170105T013443.SAFE

Identifies a Level-1C product acquired by Sentinel-2A on the 5th of January, 2017 at 1:34:42 AM. It was acquired over Tile 53NMJ(2) during Relative Orbit 031, and processed with PDGS Processing Baseline 02.04.

In addtion to the above changes, a a TCI (True Colour Image) in JPEG2000 format is included within the Tile folder of Level-1C products in this format. For more information on the TCI, see the Definitions page here.

(1) Reprocessed products - including those Reprocessed after the 6th of December - still use the old format Naming Convention (see below).

(2) The Tile Naming Convention is described in the 'Tiles and UTM Tiled Grid' section of the [PSD]. The Tiling Grid (in KML format) is availablehere.

Old format Naming Convention for Sentinel-2 Level-1C products generated before the 6th of December, 2016:

The top-level SENTINEL-2 product folder name is composed of upper-case alphanumeric characters separated by an underscore (_).

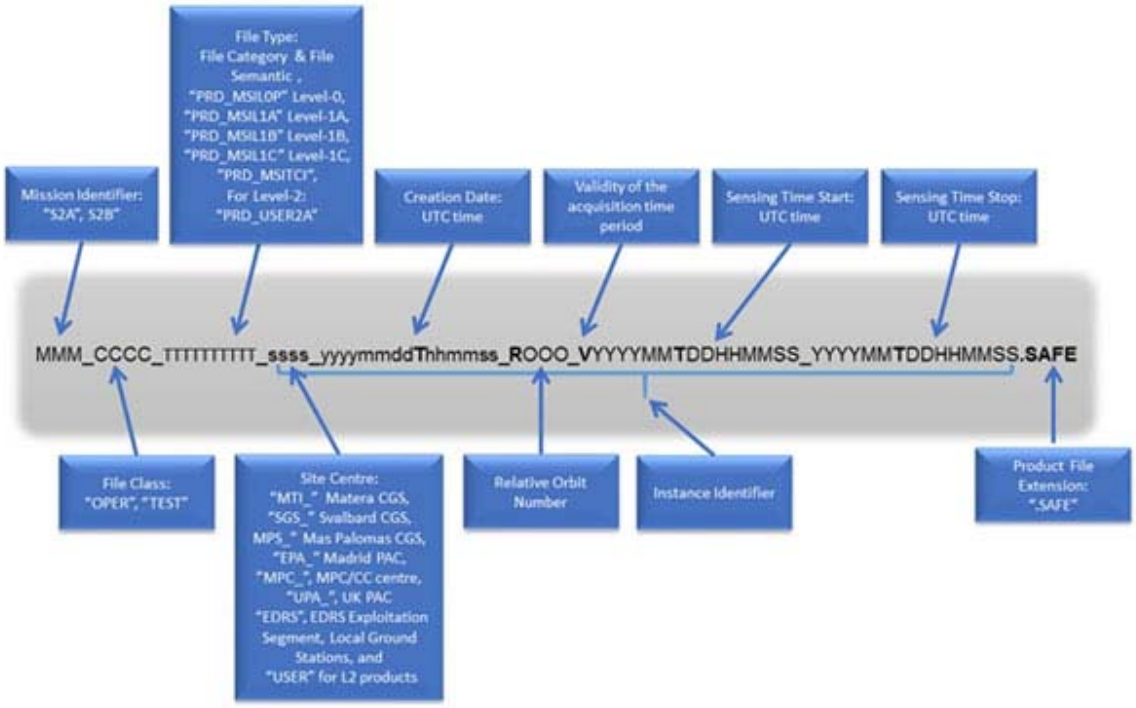

Figure 1: Product naming convention

The Mission Identifier (MMM) denotes the satellite and will be either S2A for the SENTINEL-2A instrument or S2B for the SENTINEL-2B instrument.

The File Class (CCCC) may be set to "OPER" for all products generated during the Operations Phase. During Validation or internal testing, other values can be defined.

The File Type (TTTTTTTTTT) consist of the File Category and File Semantic

TTTTTTTTTT = FFFFDDDDDD where:

FFFF = File Category and DDDDDD = Semantic Descriptor.

The File Category comprises 3 characters and an ending underscore "_". This sub-field allows the definition of file groups characterised by related information / configuration information / generated data / usage of the data / etc. The Semantic Descriptor comprises 6 characters. The Semantic Descriptor can be composed of uppercase letters, digits or underscores. The File Type for a product may be "PRD_MSIL0P" for Level-0, "PRD_MSIL1A" for Level-1A, "PRD_MSIL1B" for Level-1B, "PRD_MSIL1C" for Level-1C, "PRD_MSITCI", or for Level-2A; "PRD_USER2A"

The Instance Identifier contains the Site Centre (ssss) and Creation Date in UTC time (yyyymmddThhmmss). The Site Centre denotes the file originator or processing facility. The Site Centre may be set to "MTI_" for Matera CGS, "SGS_" for Svalbard CGS, "MPS_" for Maspalomas CGS, "EPA_" for Madrid PAC, "MPC_" for the MPC/CC centre, "UPA_", for the UK PAC, "EDRS" for the EDRS Exploitation Segment, plus Local Ground Stations ending in "L", and

"USER" for L2 products. A combination of additional fields may be included and are distinguished by the leading character:

The Relative Orbit Number (**R**OOO) where OOO represents the relative orbit number.

The Applicability Time Period: (**V**yyyymmddThhmmss_YYYYMMDDTHHMMSS) appends the Validity Period Time fields (Start and Stop).

The Product Format (SAFE) indicates that the product is in SENTINEL SAFE format

Granule naming convention:

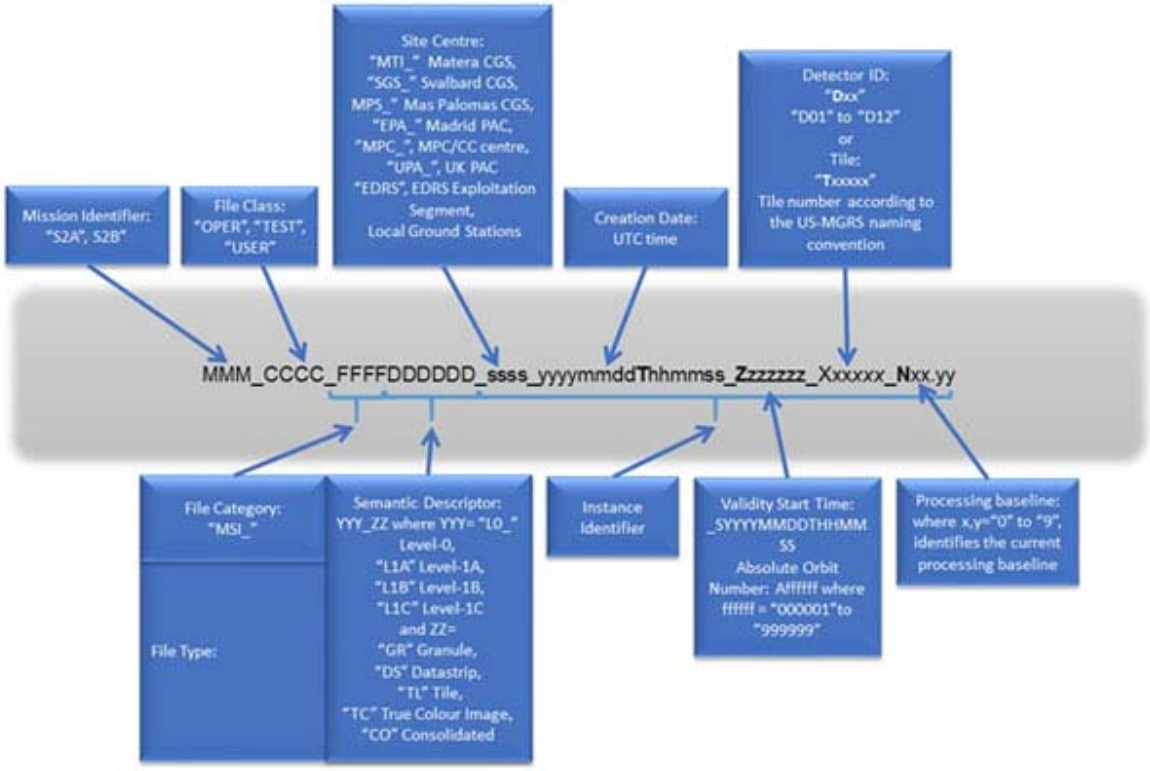

Figure 2: Granule (Granule and Tile) naming convention

Mission Identifier (MMM) and File Class (CCCCC) are as per the product naming convention.

The File Type (TTTTTTTTTT) consist of the File Category and File Semantic

TTTTTTTTTT = FFFFDDDDDD where:

FFFF = File Category and DDDDDD = Semantic Descriptor.

The File Category comprises 3 characters and an ending underscore "_". For Granules this value is set to "MSI_". The Semantic Descriptor DDDDDD = YYY_ZZ where YYY may be "L0_" for Level-0, "L1A" for Level-1A, "L1B" for Level-1B, "L1C" for Level-1C and ZZ may be "GR" for Granule, "DS" for Datastrip, "TL" for Tile, "TC" for True Colour Image, or "CO" for Consolidated.

The components of the Instance Identifier vary slightly when compared to the Product Naming Convention and may include:

The Applicability Start time (**_Z**zzzzz) = **_S**YYYYMMDDTHHMMSS Appends the Validity Start Time (granule)

or

The Absolute Orbit Number (**_Z**zzzzz) = **_A**ffffff where ffffff is the Absolute Orbit Number (tile)

The Detector ID (**_X**xxxxx) = **D**xx where xx ranges from 01 to 12 (granule)

or

Tile Number (**_X**xxxxx) = **_T**xxxxx which represents the Tile number according to the US-MGRS naming convention (tile)

The Processing Baseline Number (Nxx.yy) identifies the current processing baseline where x and y may be a digit ranging from 0 to 9.

File naming:

In addition to the fields above, filenames can include:

Band Index ID (Bxx) where xx is the band number

Completeness ID (Wx) where x = F for Full orbit x = P for Partial orbit

Degradation ID (Ly) where y = N for Nominal data y = D for Degraded data

Plus file extensions of .jp2 for JPEG2000 format image files, xml for metadata files, and .gml or Geographic Markup Language files

True Colour Images (TCI)

Following the introduction of the Compact Naming Convention on the 6th of December 2016 (see the Sentinel-2 User Guide Product Naming Convention page for details), a TCI (True Colour Image) in JPEG2000 format is included within the Tile folder of Level-1C products in this format[1]. The TCI is an RGB image built from the B02 (Blue), B03

(Green), and B04 (Red) Bands. The reflectances are coded between 1 and 255, 0 being reserved for 'No Data'. The saturation level of 255 digital counts correspond to a level of 2000 for the individual bands (i.e. reflectance of 0.2).